Available online at www.sciencedirect.com
ScienceDirect
Neurocomputing 71 (2007) 197–220
NEUROCOMPUTING
www.elsevier.com/locate/neucom

Boundedness of the nominal coefficients in Gaussian RBF neural networks

Juan Ignacio Mulero-Martínez

Departamento de Ingeniería de Sistemas y Automática, Universidad Politécnica de Cartagena, Campus Muralla del Mar, Cartagena 30203, Spain

Received 29 May 2006; received in revised form 10 January 2007; accepted 19 January 2007
Communicated by K. Li
Available online 27 April 2007

Abstract

The paper analyzes the boundedness of the coefficients involved in Gaussian expansion series. These series arise from the reconstruction of bandlimited functions, applying the sampling theorem with Gaussians as reconstruction filters. The boundedness of the ideal coefficients is a previous requirement that should be imposed to the approximation function. This is due to the fact that the coefficient sequence should be absolutely summable. With this sort of requirements, the targeted function is guaranteed to exhibit finite energy so that it will be manageable from the viewpoint of the approximation theory. On the other hand, the bounds of the coefficients affect considerably to the approximation errors and consequently to the accuracy of the estimation. The major result of this work is formalized in a series of propositions where it is stated how the coefficients are upper bounded by a signal ‘‘sinus cardinalis’’ (sinc). Finally, an energy measure of the approximation error is determined as a mean square error. In this line, a number of results are presented in both the univariate and the multivariate case showing how these errors strongly depend on the coefficients in the Gaussian expansion.

Keywords: Best approximation; Truncation errors; Sampling theory; Radial basis functions; Bandlimited functions; Existence and uniqueness; Curse of dimensionality

1. Introduction

During the past two centuries many advances in mathematical analysis have led to discover a number of important classes of approximating functions ranging from the classical Lagrange polynomial, trigonometric series and orthogonal functions to modern techniques such as radial basis functions (RBFs), multilayer feedforward neural networks (MFNNs), splines and fuzzy systems. All these approaches, under certain conditions, share the property of exhibiting function approximation abilities and have their own characteristic theory and history. A method can be considered preferable as compared to others for a given approximation problem and under some circumstances. However, each method has its pros and cons, so that it should not be asserted that an approach is the panacea in the sense of being absolutely the best. The following features of neural networks make them particularly attractive and promising for applications to modelling and control of nonlinear systems: (i) universal approximation abilities, (ii) parallel distributed processing abilities, (iii) learning and adaptation and (iv) natural fault tolerance and feasibility for hardware implementation [10,21]. The success of neural networks [3,4] is due to their learning ability and universal approximating power. The choice of the networks depends on the conditions and prior knowledge for the studied systems. If dimensionality of the input vector is not very high and the ranges of network input signals can be determined/guaranteed a priori, RBF networks shall be used to simplify the design and analysis. Due to their functional approximation capabilities, RBF networks have been seen as a good solution to interpolation problems. RBF networks are constructed from a set of nonlinear functions that are assembled into one function that can partition the search space successfully. The RBF networks have some useful properties which render them

E-mail address: juan.mulero@upct.es

0925-2312/$ - see front matter
doi:10.1016/j.neucom.2007.01.011

suitable for modelling and control. First, they belong to a class of linearly parametrized networks where the network output is connected to the tunable weights in a linear manner. The RBF uses hyper-ellipsoids to partition the pattern space. This is different from MLP networks which build its classifications from hyper-planes, defined by a weighted sum. The RBF network also requires less computation time for learning [16] and has a more compact topology [13]. Furthermore, on-line learning rules can be used to update the weights and the convergence results can be derived. The data are guaranteed to be fitted exactly if there is a basis function for each input. However, if the input data are noisy, then the noise will be learned, causing problems with generalization. Reducing the number of basis functions until an acceptable fit is achieved can reduce this problem.

Apart from pattern recognition, RBF networks have been used for a wide range of applications such as time series prediction [18], speech recognition [17] and adaptive control [19]. The use of function approximators within adaptive control designs has been popular since the publication of the papers [22]. The idea is to use standard adaptive designs for problems which depend on nonparameterized uncertainties, by utilizing function approximators to induce an appropriate approximate parameterization of unknown system functions. Recently, increasing attention has been paid to the use of artificial neural networks in nonlinear control [5,11]. In [9,19], RBFs were used for the adaptive control of SISO systems. Because RBFs are linearly parametrized, convergence results can be rigorously established.

The RBF network's advantage is that once the basis functions have all been chosen, the designer only needs to choose the coefficients for each. The coefficients are added in a linear manner and thus the solution is guaranteed since there are no local minima to avoid. In this paper, it is assumed that the only information available concerning the target function is the smoothness specified by the bandwidth, and an upper bound on the magnitude of its spectrum. No other prior information about the exact values of this function or its spectrum are known. As a result, the actual coefficient cannot be specified "a priori". Although it is difficult to determine the exact value of the coefficients, it is possible to obtain an upper bound for them. This paper focuses on the boundedness aspects of the coefficients appearing in Gaussian expansion series, leading to a wider and deeper understanding of their influences, such as those on the energy attached to the Gaussian approximation error. The main purpose is to consider the coefficients from the general point of view and to elucidate their role in applied design by using an efficient approach based on introducing the multivariate sampling theorem.

At this point, a number of questions should arise spontaneously: Why to make so efforts to derive bounds on the ideal coefficients? Which are the implications of these bounds on the design with neural networks? These or similar questions immediately pop into the mind the first time one comes across the term "boundedness". This paper aims to provide a coherent approach to the extensive ramifications originating from these, at first glance simple starting points. To this purpose it is worth emphasizing the dependence of the functional reconstruction error $\varepsilon(x)$ (sometimes called as the NN function approximation error) on the nominal coefficients of the network, besides the number of nodes required in the expansion series. Whilst the "complexity" and the "rates of approximation" have been widely studied, little care has been carried out on the derivation of explicit bounds on the nominal coefficients providing an ideal approximation. Literature in the field of neural networks reflects this situation, it emphasizes other properties in characterizing approximations. Indeed, one of the first assumptions when designing networks is the boundedness of the approximation error by a function depending on the number of network nodes. In a sense, it is public knowledge that the error decreases as the network size increases. Some results have been developed in this line and reveals how large the number of nodes should be for a specified approximation accuracy. However, the knowledge of bounds on the nominal coefficients is crucial in many practical applications, including those of modelling and control. As for the structured network modelling, in control engineering, networks are usually used to generate input/output maps using the so-called "universal approximation" property. Highlights in this field cover the complete design cycle, while providing detailed insight into most critical design issues such as tuning of network parameters to attain a prescribed degree of accuracy. As mentioned above, the accuracy achieved by an approximation scheme strongly depends on the ideal coefficients of the network. From this, it is clear that deriving bounds on the coefficients means also to get appropriate bounds on the approximation error. By restricting our attention on the adaptive control based on neural networks, in the literature special attention has been given to estimation, in particular learning procedures. Until recently, designers have mostly focused on how to achieve the desired performance requirements, and the design pays limited attention to the precise adjustment of controller gains. All the adaptive control schemes rely on the assumption that the ideal weights are unknown and may even be nonunique. Moreover, it is assumed that the weights are bounded, with the bound known, on a compact set. When control engineers design stable controllers based on neural networks, it is useful not only to track a reference trajectory, but also to ensure that all signals in the system are bounded. In this, the weight deviations or weight estimation errors deserve careful attention since candidate Lyapunov functions are commonly formulated in terms of them. In practical applications there are often unknown disturbances or modelling errors, so that even the stability in the sense of Lyapunov (SISL) is too strong to expect in closed-loop systems [15]. In all those cases, the best to be aspired after the asymptotic stability (AS) and the SISL is the uniform ultimate boundedness. This is a more practical